

# Molecular Dynamics Simulation for Proteins : GROMACS Tutorial

분자동역학 모의실험 튜토리얼

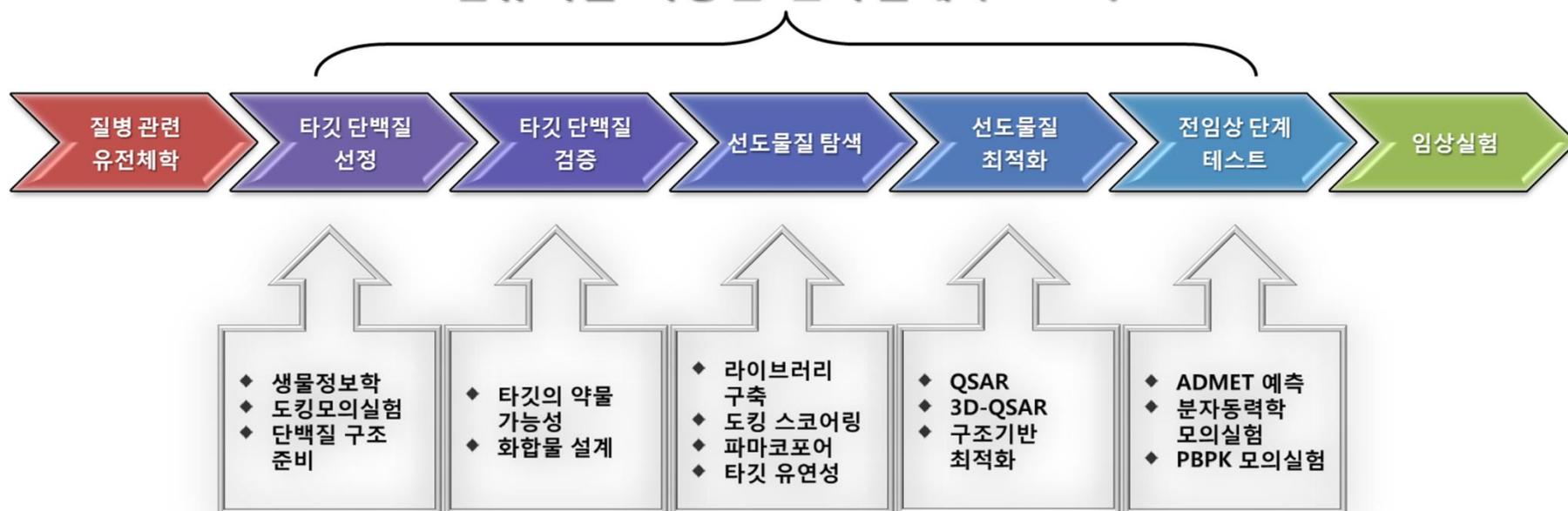
**이 윤 호**

**Yuno Lee, Ph.D**

**고등과학원 (KIAS)**

# Computer-Aided Drug Design (CADD) and Molecular Modeling

## 컴퓨터를 이용한 신약설계 (CADD)



- ❖ 합리적 분자설계를 수행하기 위해서 컴퓨터를 이용한 신약설계 방법 (CADD)이 사용되고 있으며, 이러한 *in silico* 접근법은 리간드기반 신약설계 (ligand-based drug design)와 구조기반 신약설계 (structure-based drug design)으로 구분된다.

# Molecular Dynamics (MD) simulation in CADD

---

## Molecular modeling study on Target protein

- **Molecular dynamics (MD) simulation study** for mechanism of target protein and for further finding multi-step targeting drug.
- **Molecular docking calculation** using known inhibitors to understand reaction mechanism of the protein.



## Computer-Aided Drug Design (CADD) for target protein inhibitor

- **Ligand-based and receptor-based pharmacophore modeling** for target protein.
- **Virtual screening** to find target inhibitor candidate molecules.



## Refinement and optimization of new inhibitor

- **ADME/Tox and molecular docking** to refine the candidate molecules.
- **MD simulation study** of target protein with inhibitors to optimize the structure.



Identification of **target drug candidates**

# 분자동역학 모의실험을 통한 단백질의 실시간 움직임 관찰

## 분자동역학 모의실험이란?

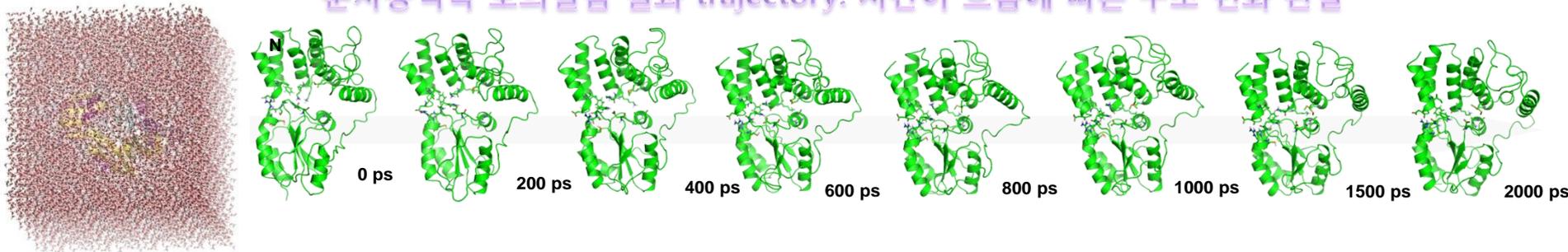
물리화학 법칙을 사용한 계산을 통하여 단백질과 같은 분자들의 실시간 움직임을 자세하게 관찰할 수 있게 해주는 이론적인 실험입니다.

## 모의실험을 위해 사용되는 슈퍼컴퓨터

슈퍼컴퓨터는 과학기술연산에 사용되는 초고속 컴퓨터입니다. 이러한 컴퓨터의 발달이 복잡한 단백질 구조의 분자동역학 모의실험을 가능하게 했습니다.



분자동역학 모의실험 결과 trajectory: 시간이 흐름에 따른 구조 변화 관찰



- ❖ MD는 미시적인 세계를 해석하는데 있어서의 이론의 한계, 실험의 한계를 극복하기 위해 사용되는 방법이다.
- ❖ 이러한 계산을 통해서 분자시스템의 화학반응 메커니즘 분석 및 시간에 따른 물성을 조사하는 것을 목표로 한다.
- ❖ 실험을 하는 것이 힘들거나 불가능한 경우 system의 성질을 알기 위해서 (실험적으로 접근 불가능한 조건이나 물질도 구현할 수 있으며, 관측하기 어려운 property들도 얻을 수 있다.)

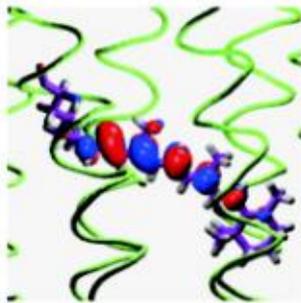
# 컴퓨터 모의실험 분야

## A. Dynamics Simulations(동역학 모의실험):

1. Molecular Dynamics (MD) Simulation:  
CHARMM, AMBER, GROMACS, NWChem, NAMD, DISCOVER,...
2. Langevin Dynamics (LD) Simulation:  
CHARMM, NAMD
3. Brownian Dynamics (BD) Simulation:  
UHBD (University of Houston BD)

### Three-Layers of Biomolecular Simulation

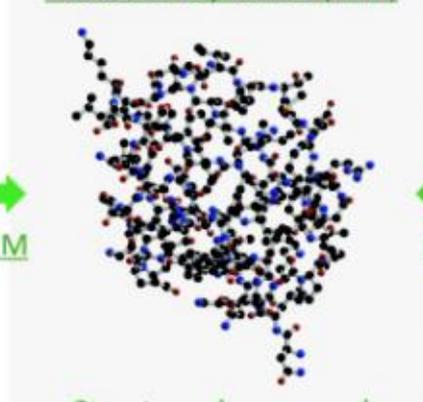
#### Quantum Chemical (QM)



Coupling between  
reactions and motions

QM/MM

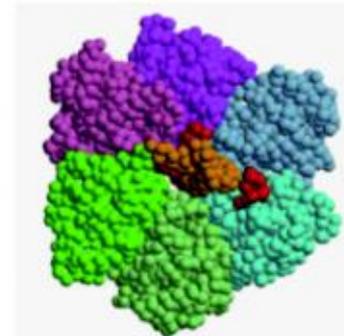
#### Molecular Dynamics (MM)



Structure changes and  
molecular interactions

MM/CG

#### Coarse-Grained Model (CG)



Large spatial/temporal scale  
events in cellular context

# GROMACS: FAST, FLEXIBLE, FREE

---

- ❖ 분자동역학 모의실험을 하기 위한 도구들로 CHARMM, NAMD, Discover, **GROMACS**, 등 여러가지 프로그램들이 있지만, 그 중에서도 **GROMACS**는 GNU정신에 입각하여 공짜로 다운받을 수 있고, 계산이 빠르고, 다양한 분석이 가능합니다.
- ❖ 또, 수 백에서 수백만의 입자를 갖는 시스템들에서 뉴턴 운동방정식을 계산할 수 있을 정도로 분자동역학을 수행하는데 있어 다재다능한 패키지입니다.
- ❖ 이 소프트웨어는 최초로 복잡한 공유결합 상호작용들을 가진 지질과 단백질과 같은 생화학적 분자들을 위해 개발되었지만, 비공유결합 상호작용들을 계산하는데 있어 매우 빠르기 때문에 많은 그룹들이 비생물학적 시스템들에서도 사용하고 있습니다. 예. polyemrs
- ❖ **GROMACS**는 현대 분자동역학 도구에서 기대되는 모든 유용한 알고리즘들을 지원할 뿐만 아니라 몇몇 매우 특별한 점들도 가지고 있어 다른 프로그램들과 구별됩니다.

# The global MD algorithm

## 1. Input initial conditions

Potential interaction  $V$  as a function of atom positions  
 Positions  $\mathbf{r}$  of all atoms in the system  
 Velocities  $\mathbf{v}$  of all atoms in the system



repeat 2,3,4 for the required number of steps:

## 2. Compute forces

The force on any atom

$$\mathbf{F}_i = -\frac{\partial V}{\partial \mathbf{r}_i}$$

is computed by calculating the force between non-bonded atom pairs:

$$\mathbf{F}_i = \sum_j \mathbf{F}_{ij}$$

plus the forces due to bonded interactions (which may depend on 1, 2, 3, or 4 atoms), plus restraining and/or external forces.

The potential and kinetic energies and the pressure tensor are computed.



## 3. Update configuration

The movement of the atoms is simulated by numerically solving Newton's equations of motion

$$\frac{d^2 \mathbf{r}_i}{dt^2} = \frac{\mathbf{F}_i}{m_i}$$

or

$$\frac{d\mathbf{r}_i}{dt} = \mathbf{v}_i; \quad \frac{d\mathbf{v}_i}{dt} = \frac{\mathbf{F}_i}{m_i}$$



## 4. if required: Output step

write positions, velocities, energies, temperature, pressure, etc.

## Energy terms of force field

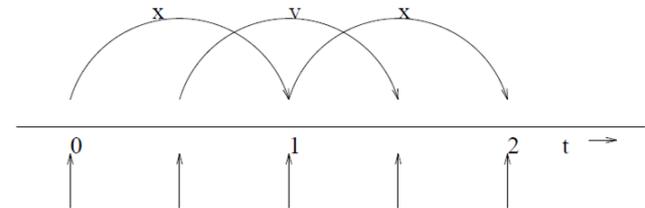
$$E(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \sum_{bonds} \frac{1}{2} k_i^b (d_i - d_i^0)^2$$

$$+ \sum_{angles} \frac{1}{2} k_i^\theta (\theta_i - \theta_i^0)^2$$

$$+ \sum_{torsions} k_i^\phi [1 + \cos(n_i \phi_i - \delta_i)]$$

$$+ \frac{1}{2} \sum_{nonbond} \left( \epsilon_{ij}^{min} \left[ \left( \frac{d_{ij}^{min}}{d_{ij}} \right)^{12} - 2 \left( \frac{d_{ij}^{min}}{d_{ij}} \right)^6 \right] + \frac{q_i q_j}{\epsilon d_{ij}} \right)$$

## Integrating the equation of motion



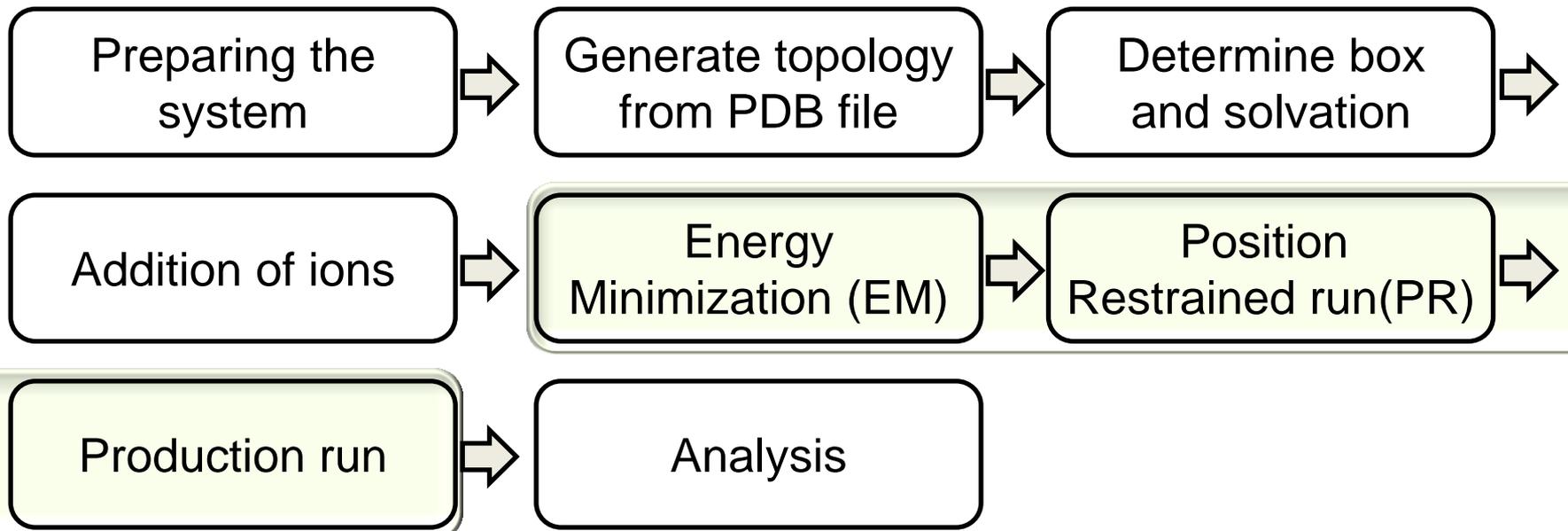
The **Leap-Frog integration** method. The algorithm is called Leap-Frog because  $\mathbf{r}$  and  $\mathbf{v}$  are leaping like frogs over each other's backs.

$$\mathbf{v}(t + \frac{1}{2} \Delta t) = \mathbf{v}(t - \frac{1}{2} \Delta t) + \frac{\Delta t}{m} \mathbf{F}(t)$$

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \Delta t \mathbf{v}(t + \frac{1}{2} \Delta t)$$

# Molecular Dynamics (MD) simulation

## Flow chart of MD simulation procedures



3 MD processes

GROMACS 설치관련:

[http://newton.kias.re.kr/~yunolee1/homepage/gromacs\\_install.html](http://newton.kias.re.kr/~yunolee1/homepage/gromacs_install.html)

Yuno's home - Chrome  
 newton.kias.re.kr/~yunolee1/homepage/gromacs\_install.html

Liposome [Drug Design] [Dynamics] [Physics] [Math] 통계역학 프로그래밍 [MATLAB] [생물정보학] [Systems biology] [생화학] Quantum [Statistics]

KOREA INSTITUTE FOR ADVANCED STUDY | COMPUTATIONAL SCIENCES

# Yuno Lee's Home

Profile  
 Programing lecture  
 Graphic lecture  
 Home page lecture  
 English lecture  
 Research  
 Favorite  
 Think

## 우분투에서 GROMACS 5.0.2 설치하기

글쓴이 : 이윤호 (고등과학원 연구원), 문의: yunolee1@gmail.com; yunolee1@kias.re.kr

먼저, openmpi 설치를 하신 후, GROMACS 홈페이지에 나와 있는 방법으로 설치하시면 됩니다.

**openmpi 설치**  
<http://www.sysads.co.uk/2014/05/install-open-mpi-1-8-ubuntu-14-04-13-10/>  
<http://www.open-mpi.org/>

- 다운로드 및 infiniband headers in the standard package libibnetdisc-dev 설치  
`wget https://www.open-mpi.org/software/ompi/v1.8/downloads/openmpi-1.8.3.tar.gz`  
`sudo apt-get install libibnetdisc-dev`
- 압축풀기 (Decompress the downloaded file)  
`tar -xvf openmpi-1.8.3.tar.gz`  
`cd openmpi-1.8.3`
- 설치 위치 설정하기 (Configure the installation file)  
`./configure --prefix="/home/$USER/.openmpi"`
- 설치 (Install OpenMPI (This path will take time to complete))  
`make`  
`sudo make install`
- 환경적용 (Setup path in Environment Variable)  
`export PATH="$PATH:/home/$USER/.openmpi/bin"`  
`export LD_LIBRARY_PATH="$LD_LIBRARY_PATH:/home/$USER/.openmpi/lib/"`
- 테스트 (Test if install was successful)  
`mpirun`

=====

**GROMACS 홈페이지 나와있는 설치순서**  
[http://www.gromacs.org/Documentation/Installation\\_Instructions](http://www.gromacs.org/Documentation/Installation_Instructions)

- 1) Get the latest version of your C and C++ compilers.
- 2) Check that you have CMake version 2.8.8 or later.
- 3) Get and unpack the latest version of the GROMACS tarball.
- 4) Make a separate build directory and change to it.
- 5) Run cmake with the path to the source as an argument
- 6) Run `make`, `make check`, and `make install`.

GROMACS 설치관련:

[http://newton.kias.re.kr/~yunolee1/homepage/gromacs\\_install.html](http://newton.kias.re.kr/~yunolee1/homepage/gromacs_install.html)

cd gromacs-5.0.2

# http://cafe.naver.com/insilico (네이버 카페 인실리코)

분자모델링 연구 Cafe

경상대학교 생화학과  
생물정보학 연구실  
Gyeongsang National University (GNU)  
Bioinformatics Lab (http://bio.gnu.ac.kr)

## InSilico



http://cafe.naver.com/insilico

카페정보 - 나의활동 -

매니저 이윤호 (youknow1...)  
업버수 222  
프로필 since 2007.05.14

> 초대하기 > 채팅하기

카페 글쓰기

☞ 관리 카페 통계

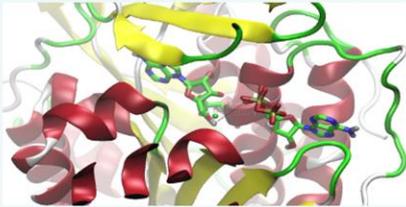
### Welcome

분자모델링카페 InSilico 입니다

Chief Manager 이윤호  
Assistant Manager 김성미  
Fine Manager 백아영  
Members List 01이윤호, 02황성환, 03김성미, 04손민경,  
05백아영, 06박창인, 07손영식

Since 2007.08.23

Supervisor 이근우 교수님



전체글보기 (1,489)

- 카페로그보기
- 베스트게시물
- 카페 캘린더
- 미투데이보기

카페북 책꽂이

인실리코

- 알립니다
- 가입인사
- 등업 요청
- 추천게시판
- 유머게시판
- 자유게시판
- 메모게시판
- 출석체크
- Q & A

계산생물학

- 소개
- 분자모델링
- 단백질공학
- 분자동역학
- 단백질구조 예측

생물정보학

신약설계

맞춤의학

특허관련

**알립니다** 더보기 >

공지 해당조건을 넘을 경우에 ..	0
공지 2010년 이월 회비 관련.. [5]	14
공지 청소 담당 구역입니다~ [3]	11
공지 카페에 가입 하실때, .. [2]	46
공지 인실리코 [in silico] 사진..	47
• 2010년 이월 회비 관련해.. [5]	14
• [커피미팅]김성미씨편	12
• [커피미팅] 황성환씨편 HIV~.. [3]	3
• 청소 담당 구역입니다~ [3]	11
• [논문스크랩] R, 회귀.. [2]	16
• [2009년 04월 06일 ~ 2009년..]	10
• [2009년 03월 30일 ~ 2009년..]	4

**전체글보기** 더보기 >

공지 해당조건을 넘을 경우에 자동으로 등업..	이윤호	2010/01/19	0
공지 2010년 이월 회비 관련해서 말씀드.. [5]	공정희	2010/01/05	14
공지 청소 담당 구역입니다~ [3]	식이입니다	2009/12/21	11
공지 카페에 가입 하실때, 가입 질문에 .. [2]	미야	2008/12/04	46
공지 인실리코 [in silico] 사진적 의미	황성환	2008/10/31	47
• 계산생물학의 소개 (Computational Bl..	이윤호	2010/01/20	2

**메모게시판** 더보기 >

허걱 정말이네... 민경이는 일..	2010.01.19
어렸~ 2010년 되서 글 올린 사람..	2010.01.17
프린트 문제 외국인들한테 1차 ..	2008.12.22
연구에 몰입하기!	2008.12.07
자 다들 새로운 마음으로 연구해..	2008.11.16
다들 나 없어도 잘 지내시고, 한..	2008.08.11

6 경상남도진주



COUNTER

TODAY 110

TOTAL 95168

환율조회 | 환율기상 | 외환은행기준

	기준	실세	말세
1달권	1131.90	1151.70	1112.10
100엔	1241.80	1263.53	1220.07
1유로	1609.34	1641.36	1577.32

**주식**

주가검색  검색

코스피 1,715.35 (▲5.13)

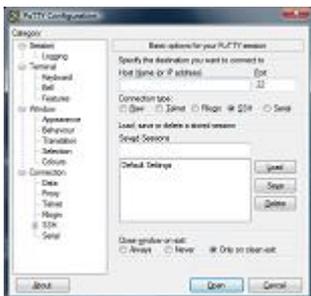
코스닥 546.60 (▼2.54)

• [오늘의포인트] 가치투..

• 공모주 시장 탐방담..

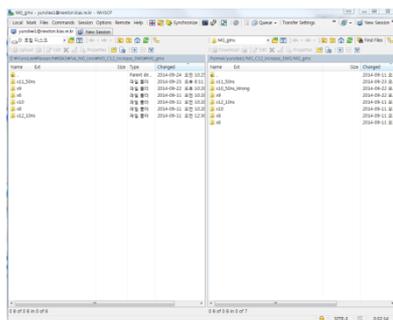
업데이트 시각 01/20 12:23

# SSH / SCP (Windows OS)



## Download PuTTY

- ❖ PuTTY is an SSH and telnet client, developed originally by Simon Tatham for the Windows platform. PuTTY is open source software that is available with source code and is developed and supported by a group of volunteers.
- ❖ You can download PuTTY:  
<http://www.chiark.greenend.org.uk/~sgtatham/putty/download.html>

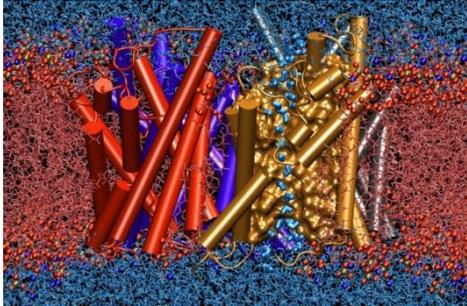


## Download WinSCP

- ❖ WinSCP is an open source free SFTP client, FTP client, WebDAV client and SCP client for Windows. Its main function is file transfer between a local and a remote computer. Beyond this, WinSCP offers scripting and basic file manager functionality.
- ❖ You can download WinSCP:  
<http://winscp.net/eng/download.php>

# Visualizations – VMD, PyMOL

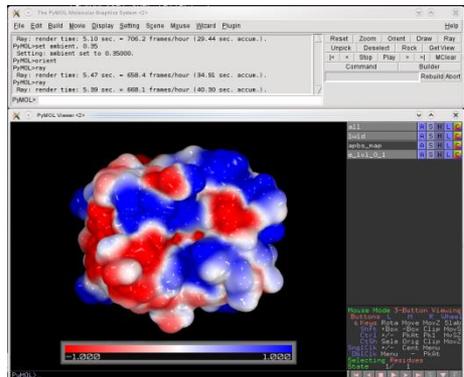
## Download VMD



- ❖ VMD is a molecular visualization program for displaying, animating, and analyzing large biomolecular systems using 3-D graphics and built-in scripting. VMD supports computers running MacOS X, Unix, or Windows, is distributed free of charge, and includes source code.

- ❖ You can download VMD:  
<http://www.ks.uiuc.edu/Development/Download/download.cgi?PackageName=VMD>

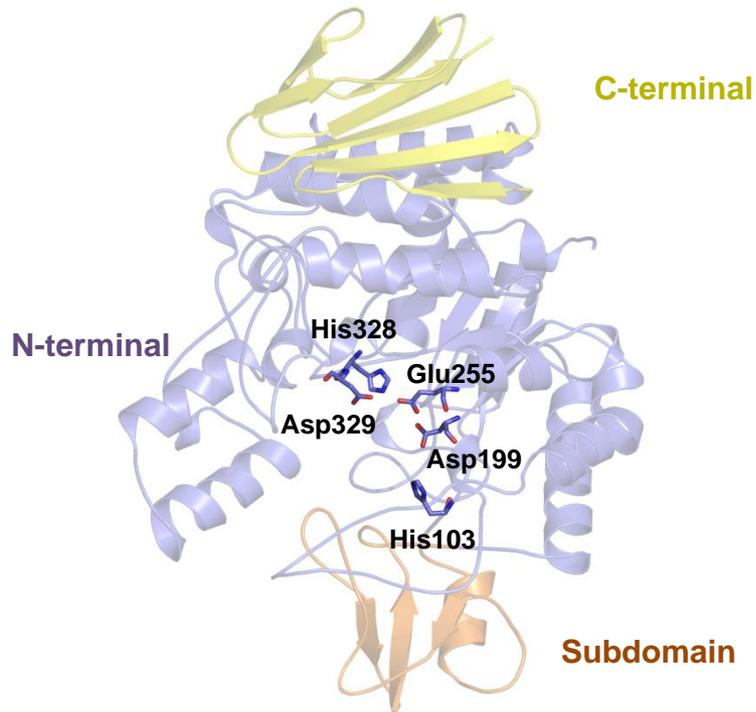
## Download PyMOL



- ❖ PyMOL is a user-sponsored molecular visualization system on an open-source foundation, maintained and distributed by Schrödinger.
- ❖ You can download PyMOL:  
<http://pymol.org/educational/>

# Preparing the system: Crystal structure of $\alpha$ -glucosidase from *B. cereus* (PDBID: 1UOK)

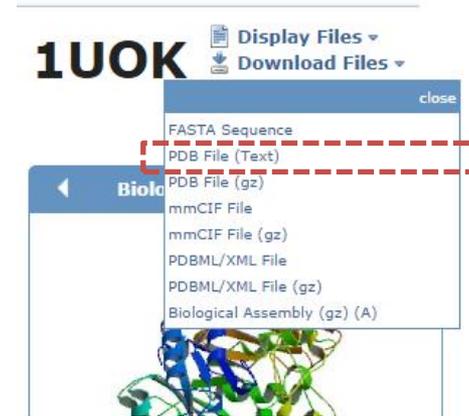
## Crystal structure of *B. cereus* $\alpha$ -glucosidase



**PDB ID:** 1UOK (*B. cereus*)

**Resolution:** 2.0 Å

“Catalytic triad residues: Asp214, Glu276, Asp349”



- ❖ Download the PDB 1UOK from RCSB website(<http://www.rcsb.org>)
- ❖ Delete crystal water molecules using plain text editor such as vi, emacs for Linux; and AcroEdit for Windows.
- ❖ Check missing residues in your pdb file

# Generate topology from PDB file

```
pdb2gmx -f pro.pdb -o pro.gro -p pro.top -i pro.itp -ff amber03 -water spce -ighn
```

## Input file

pro.pdb = the modified pdb file (1UOK).

## Output files

pro.gro = a post-processed structure file.

pro.top = the topology for the molecule.

pro.itp = a position restraint file.

Options -ff: force field selection.

-ighn: Ignore H atoms in the PDB file.

Total charge -39.000 e

## pro.gro

OLIGO-1,6-GLUCOSIDASE

9068

1MET	N	1	4.039	4.973	-0.502
1MET	H1	2	4.106	5.047	-0.499
1MET	H2	3	4.015	4.946	-0.408

## pro.top

[ molecules ]

; Compound	#mols
Protein_chain_A	1

## pro.itp

[ position\_restraints ]

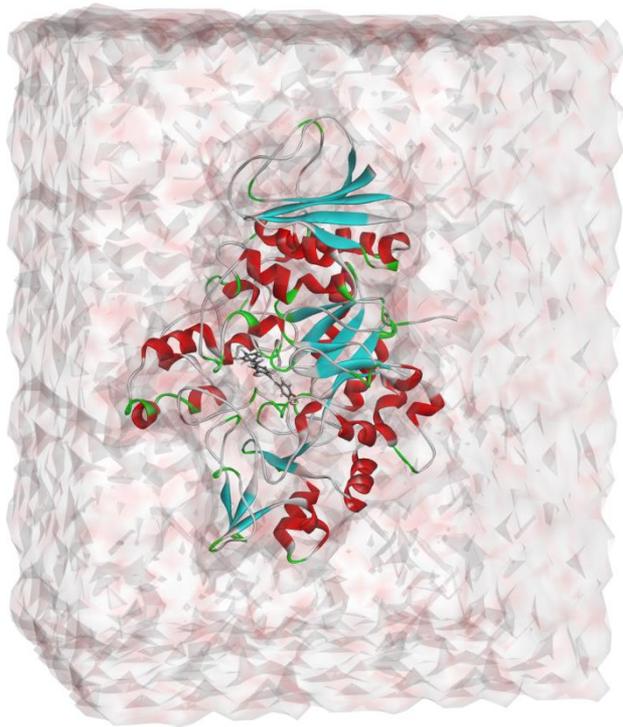
; atom	type	fx	fy	fz
1	1	1000	1000	1000
5	1	1000	1000	1000
7	1	1000	1000	1000

# Determine box and solvation

```
editconf -f pro.gro -o prob.gro -d 1.2
```

Option `-bt`: box type (triclinic, cubic, dodecahedron)

```
genbox -cp prob.gro -cs spc216.gro -o b4em.gro -p pro.top
```



Orthorhombic box (1.2 nm from protein edges)

After `editconf` command,

```
Read 9068 atoms
Volume: 1169.88 nm^3, corresponds to roughly 526400
electrons
No velocities found
  system size :  5.835  7.565  6.374 (nm)
   center     :  5.529  4.228  1.140 (nm)
  box vectors : 10.610 10.610 12.000 (nm)
  box angles  :  90.00  90.00 120.00 (degrees)
  box volume  :1169.88                (nm^3)
   shift     : -1.411  0.754  3.247 (nm)
 new center   :  4.117  4.982  4.387 (nm)
 new box vectors :  8.235  9.965  8.774 (nm)
 new box angles :  90.00  90.00  90.00 (degrees)
 new box volume :  720.01                (nm^3)
```

After `genbox` command,

```
pro.top
[ molecules ]
; Compound      #mols
Protein_chain_A  1
SOL              20930
```

# Addition of ions

Total charge -39.000 e (obtained from pdb2gmx)

```
grompp -f em1.mdp -c b4em.gro -o em_tmp.tpr -p pro.top
```

; em1.mdp - used as input into grompp to generate em\_tmp.tpr and em1.tpr

; Parameters describing what to do, when to stop and what to save

```
integrator          = steep                ; Algorithm (steep = steepest descent minimization)
emtol               = 1000.0              ; Stop minimization when the maximum force < 1000.0 kJ/mol/nm
emstep              = 0.01                ; Energy step size
nsteps              = 50000               ; Maximum number of (minimization) steps to perform
```

; Parameters describing how to find the neighbors of each atom and how to calculate the interactions

```
nstlist             = 1                  ; Frequency to update the neighbor list and long range forces
ns_type             = grid               ; Method to determine neighbor list (simple, grid)
rlist               = 1.0                ; Cut-off for making neighbor list (short range forces)
coulombtype         = PME                 ; Treatment of long range electrostatic interactions
rcoulomb            = 1.0                ; Short-range electrostatic cut-off
rvdw                = 1.0                ; Short-range Van der Waals cut-off
ppbc                 = xyz                ; Periodic Boundary Conditions (yes/no)
```

```
genion -s em_tmp.tpr -o ion_add.gro -p pro.top -np 39 -pname NA -nn 0 -nname CL
```

Or 

```
genion -s em_tmp.tpr -o ion_add.gro -p pro.top -pname NA -nname CL -neutral -conc 0.1
```

```
pro.top
[ molecules ]
; Compound      #mols
Protein_chain_A  1
SOL              20805
NA               82
CL              43
```

Option `-conc`: Specify salt concentration (mol/liter). This will add sufficient ions to reach up to the specified concentration as computed from the volume of the cell in the input `.tpr` file.

# Energy Minimization (EM)

```
grompp -f em1.mdp -c ion_add.grd -o em1.tpr -p pro.top
```

```
mdrun -v -deffnm em1
```

or

```
mpirun -np 12 mdrun -v -s em1.tpr -c after_em1.gro -deffnm em1 >& em1.job
```

## Output files

em1.log: Log file of the EM process

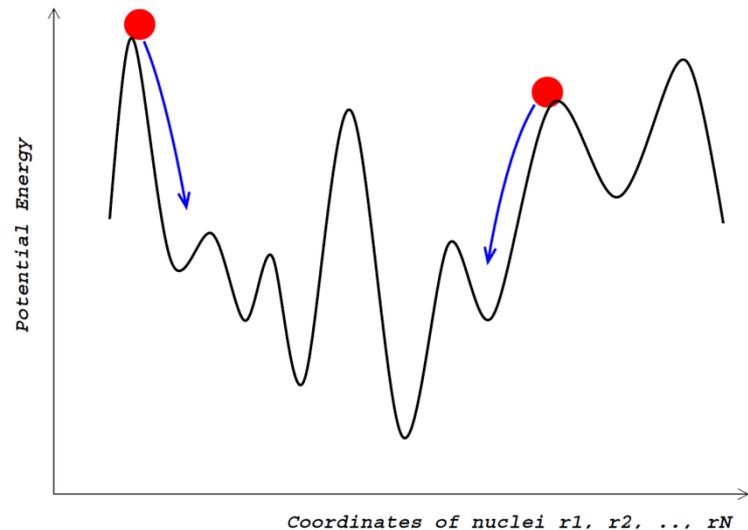
em1.edr: Binary energy file

em1.trr: Binary full-precision trajectory

em1.gro: Energy-minimized structure

## Algorithms:

Steepest descent, Conjugate gradient



# Position Restrained run(PR) - NVT

```
grompp -f pr_NVT.mdp -o pr_NVT.tpr -c after_em1.grd -p pro.top
```

```
mdrun -v -c after_pr_NVT -deffnm pr_NVT
```

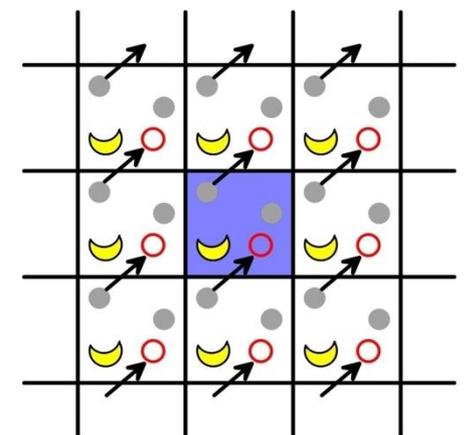
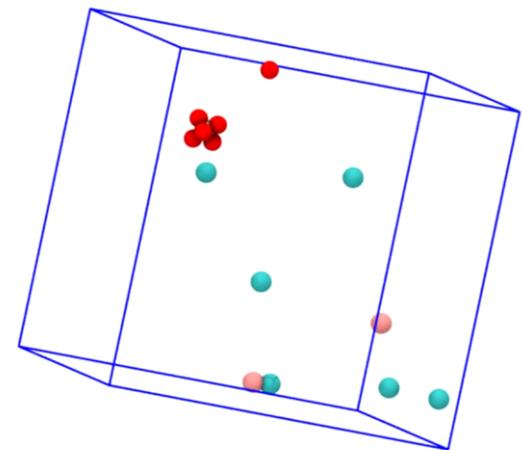
```

title                = NVT equilibration
define               = -DPOSRES                ; position restrain the protein
; Run parameters
integrator           = md                      ; leap-frog integrator
nstps                = 50000                  ; 2 * 50000 = 100 ps
dt                  = 0.002                  ; 2 fs
; Output control
nstxout             = 100                    ; save coordinates every 0.2 ps
nstvout             = 100                    ; save velocities every 0.2 ps
nstenergy           = 100                    ; save energies every 0.2 ps
nstlog              = 100                    ; update log file every 0.2 ps
; Bond parameters
continuation        = no                     ; first dynamics run
constraint_algorithm = lincs                  ; holonomic constraints
constraints          = all-bonds              ; all bonds (even heavy atom-H bonds) constrained
lincs_iter          = 1                       ; accuracy of LINCS
lincs_order         = 4                       ; also related to accuracy
; Neighborsearching
ns_type             = grid                    ; search neighboring grid cells
nstlist             = 5                       ; 10 fs
rlist               = 1.0                     ; short-range neighborlist cutoff (in nm)
rcoulomb            = 1.0                     ; short-range electrostatic cutoff (in nm)
rvdw                = 1.0                     ; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype         = PME                     ; Particle Mesh Ewald for long-range electrostatics
pme_order           = 4                       ; cubic interpolation
fourierspacing     = 0.16                    ; grid spacing for FFT
; Temperature coupling is on
tcoupl              = V-rescale                ; modified Berendsen thermostat
tc-grps             = Protein                 Non-Protein      ; two coupling groups - more accurate
tau_t               = 0.1                     0.1                ; time constant, in ps
ref_t               = 300                     300                ; reference temperature, one for each group, in K
; Pressure coupling is off
pcoupl              = no                      ; no pressure coupling in NVT
; Periodic boundary conditions
pbc                 = xyz                     ; 3-D PBC
; Dispersion correction
DispCorr            = EnerPres                ; account for cut-off vdW scheme
; Velocity generation
gen_vel             = yes                     ; assign velocities from Maxwell distribution
gen_temp            = 300                     ; temperature for Maxwell distribution
gen_seed            = -1                       ; generate a random seed

```

Canonical (NVT, Helmholtz free-energy)

$$T = \sum m \langle v^2 \rangle / (3k_b)$$



# Position Restrained run(PR) - NPT

```
grompp -f pr_NPT.mdp -o pr.tpi -c after_pr_NVT.grd -p pro.top
```

```
mdrun -v -c after_pr -deffnm pr
```

Isothermal-isobaric (NPT, Gibbs free-energy)

$P = \text{kinetic} + \text{virial contributions}$

```

title = NPT equilibration
define = -DPOSRES ; position restrain the protein
; Run parameters
integrator = md ; leap-frog integrator
nsteps = 50000 ; 2 * 50000 = 100 ps
dt = 0.002 ; 2 fs
; Output control
nstxout = 100 ; save coordinates every 0.2 ps
nstvout = 100 ; save velocities every 0.2 ps
nstenergy = 100 ; save energies every 0.2 ps
nstlog = 100 ; update log file every 0.2 ps
; Bond parameters
continuation = yes ; Restarting after NVT
constraint_algorithm = lincs ; holonomic constraints
constraints = all-bonds ; all bonds (even heavy atom-H bonds) constrained
lincs_iter = 1 ; accuracy of LINCS
lincs_order = 4 ; also related to accuracy
; Neighborssearching
ns_type = grid ; search neighboring grid cells
nstlist = 5 ; 10 fs
rlist = 1.0 ; short-range neighborlist cutoff (in nm)
rcoulomb = 1.0 ; short-range electrostatic cutoff (in nm)
rvdw = 1.0 ; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype = PME ; Particle Mesh Ewald for long-range electrostatics
pme_order = 4 ; cubic interpolation
fourierspacing = 0.16 ; grid spacing for FFT
; Temperature coupling is on
tcoupl = V-rescale ; modified Berendsen thermostat
tc-grps = Protein Non-Protein ; two coupling groups - more accurate
tau_t = 0.1 0.1 ; time constant, in ps
ref_t = 300 300 ; reference temperature, one for each group, in K
; Pressure coupling is on
pcoupl = Parrinello-Rahman ; Pressure coupling on in NPT
pcouptype = isotropic ; uniform scaling of box vectors
tau_p = 2.0 ; time constant, in ps
ref_p = 1.0 ; reference pressure, in bar
compressibility = 4.5e-5 ; isothermal compressibility of water, bar^-1
refcoord_scaling = com
; Periodic boundary conditions
pbc = xyz ; 3-D PBC
; Dispersion correction
DispCorr = EnerPres ; account for cut-off vdW scheme
; Velocity generation
gen_vel = no ; Velocity generation is off

```

# Production run

```
grompp -f md_NPT.mdp -o md1.tpr -c after_pr.gro -p pro.top
```

```
mdrun -v -deffnm md1
```

Check after grompp command

Estimate for the relative computational load of the PME mesh part: 0.25

For rectangular boxes the optimal particle-particle(PP) to PME node ratio is usually 3:1, for rhombic dodecahedra usually 2:1.

```

title = MD
; Run parameters
integrator = md ; leap-frog integrator
nsteps = 500000 ; 2 * 500000 = 1000 ps, 1 ns
dt = 0.002 ; 2 fs
; Output control
nstxout = 1000 ; save coordinates every 2 ps
nstvout = 1000 ; save velocities every 2 ps
nstxtcout = 1000 ; xtc compressed trajectory output every 2 ps
nstenergy = 1000 ; save energies every 2 ps
nstlog = 1000 ; update log file every 2 ps
; Bond parameters
continuation = yes ; Restarting after NPT
constraint_algorithm = lincs ; holonomic constraints
constraints = all-bonds ; all bonds (even heavy atom-H bonds) constrained
lincs_iter = 1 ; accuracy of LINCS
lincs_order = 4 ; also related to accuracy
; Neighborssearching
ns_type = grid ; search neighboring grid cells
nstlist = 5 ; 10 fs
rlist = 1.0 ; short-range neighborlist cutoff (in nm)
rcoulomb = 1.0 ; short-range electrostatic cutoff (in nm)
rvdw = 1.0 ; short-range van der Waals cutoff (in nm)
; Electrostatics
coulombtype = PME ; Particle Mesh Ewald for long-range electrostatics
pme_order = 4 ; cubic interpolation
fourierspacing = 0.16 ; grid spacing for FFT
; Temperature coupling is on
tcoupl = V-rescale ; modified Berendsen thermostat
tc-grps = Protein Non-Protein ; two coupling groups - more accurate
tau_t = 0.1 0.1 ; time constant, in ps
ref_t = 300 300 ; reference temperature, one for each group, in K
; Pressure coupling is on
pcoupl = Parrinello-Rahman ; Pressure coupling on in NPT
pcouptype = isotropic ; uniform scaling of box vectors
tau_p = 2.0 ; time constant, in ps
ref_p = 1.0 ; reference pressure, in bar
compressibility = 4.5e-5 ; isothermal compressibility of water, bar^-1
; Periodic boundary conditions
pbc = xyz ; 3-D PBC
; Dispersion correction
DispCorr = EnerPres ; account for cut-off vdW scheme
; Velocity generation
gen_vel = no ; Velocity generation is off

```

# Analysis

## ❖ Check trajectory using VMD program

```
make_ndx -f md1.tpr -o in_md1.ndx (except for water molecule)
```

```
editconf -f after_pr.gro -n in_md1.ndx -o after_pr.pdb (pdb file contained only protein)
```

```
trjconv -f md1.xtc -s md1.tpr -n in_md1.ndx -o trajout.xtc -pbc nojump
```

Option `-pbc nojump`: to remove PBC jumping effect in trajectory

`-dt`: Only write frame when  $t \text{ MOD } dt = \text{first time (ps)}$

it is possible to reduce the number of frames in the output.

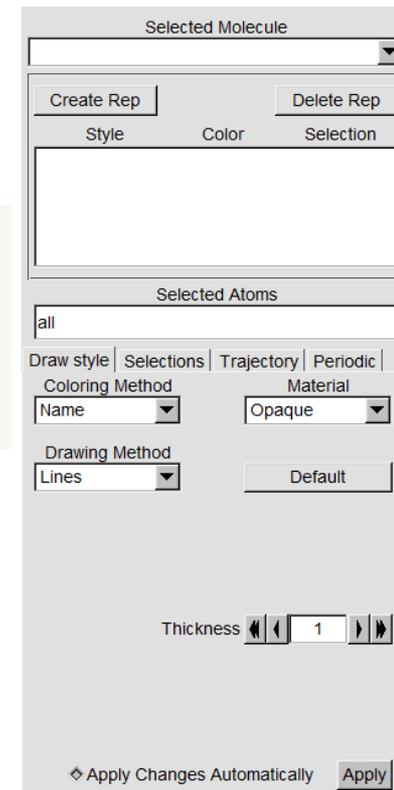
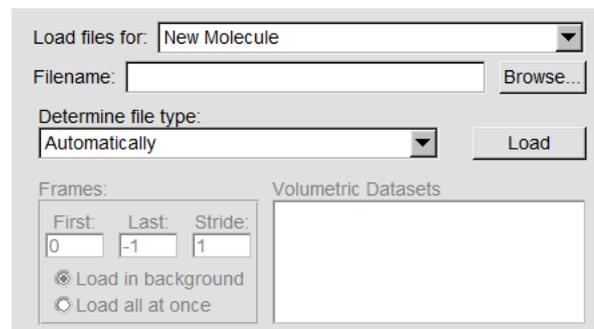
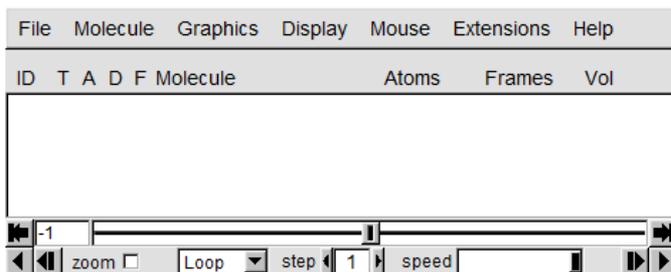
Transfer **after\_pr.pdb** and **trajout.xtc** files from Linux server to Window PC

In VMD program,

File -> New Molecule... -> Browse... -> select **after\_pr.pdb** -> Load

Select **trajout.xtc** -> Load (should keep Load files for: 0: after\_pr.pdb)

Graphics -> Representations...



# Analysis

---

## ❖ RMSD, RMSF, Radius of Gyration, Energy, Distance, H-bond Measure

### RMSD (Root Mean Square Deviation)

```
g_rms -f md1.xtc -s after_pr.pdb -o Ca_RMSD.xvg  
-> C alpha to C alpha
```

### RMSF (Root Mean Square Fluctuation)

```
g_rmsf -f md1.xtc -s md1.tpr -n in_md1.ndx -o Ca_RMSF.xvg  
-> C alpha to C alpha
```

### Radius of Gyration

```
g_gyrate -f md1.xtc -s md1.tpr -o gyrate.xvg
```

### Potential Energy Measure

```
g_energy -f md1.edr -s md1.tpr -o energy.xvg  
-> 10 0
```

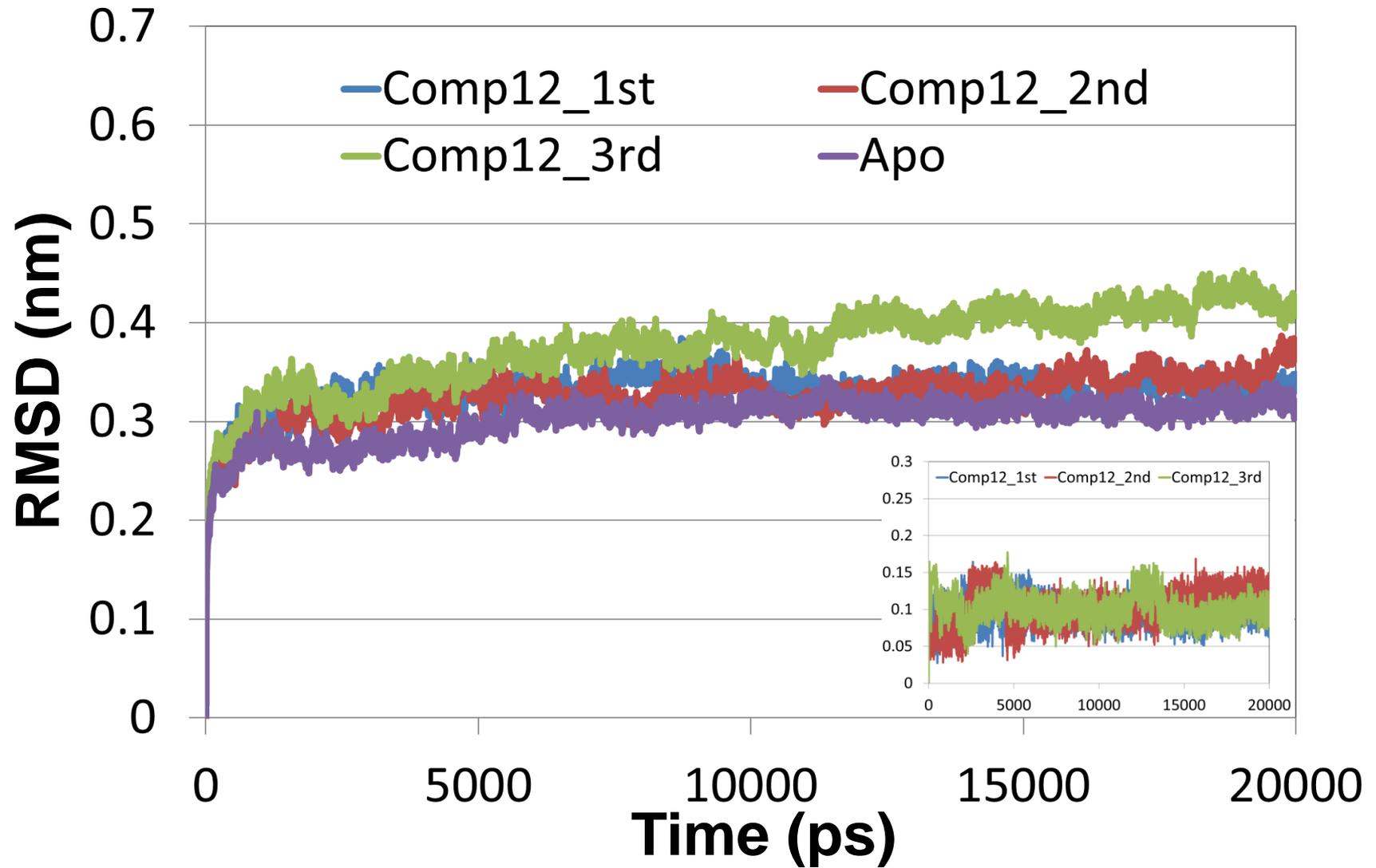
### Distance Measure

```
g_dist -f md1.xtc -s md1.tpr -n in_md1.ndx -o dist.xvg
```

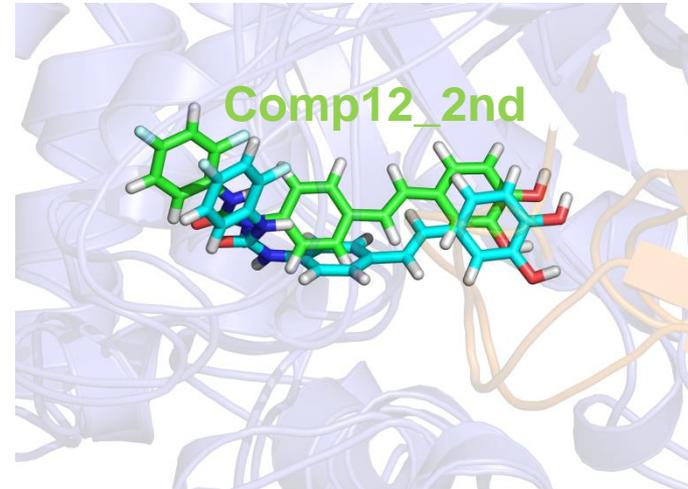
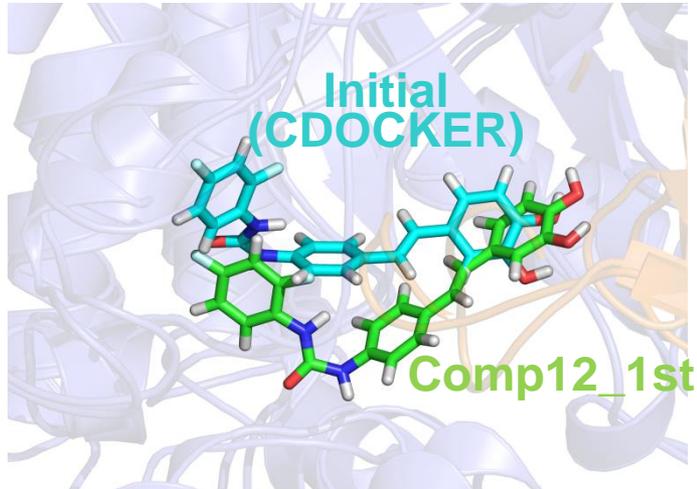
### H-bond Measure

```
g_hbond -f md1.xtc -s md1.tpr -n in_md1.ndx -num backbone_hbond.xvg  
-> backbone to backbone
```

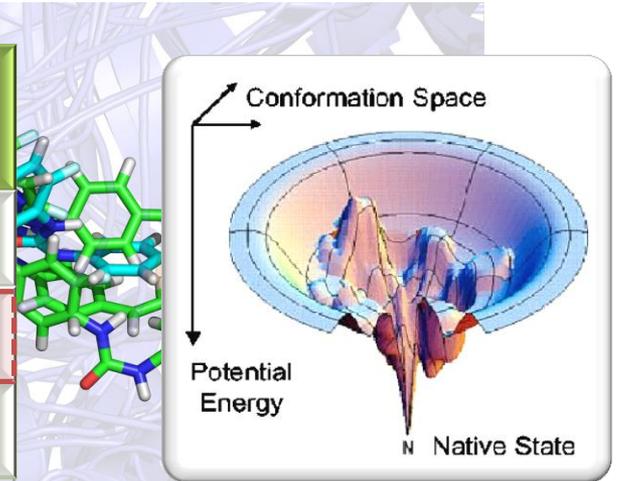
# Molecular dynamics (MD) simulation result: System stability check using RMSD plots



# MD simulation result: binding modes of **comp12** in representative structures



System	Interaction Energy (kJ/mol)	Van der Waals Energy (kJ/mol)	Electrostatic Energy (kJ/mol)
Comp12_1st	-235.876	-168.048	-67.8278
Comp12_2nd	-253.253	-202.956	-50.2964
Comp12_3rd	-203.842	-159.364	-44.4776



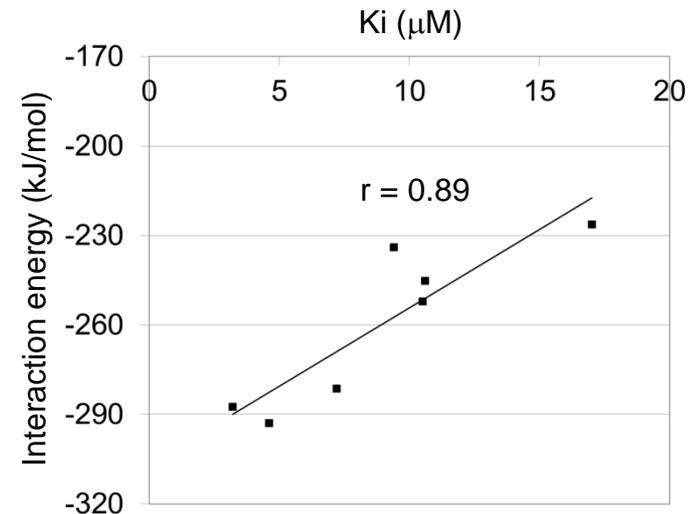
“2nd system was selected as **the lowest interaction energy structure**”

# Validation of final binding site showing linearity between **exp. Ki** and **calculated energy**

To validate the final binding mode, **exp. Ki** values were correlated with **calculated interaction energy**

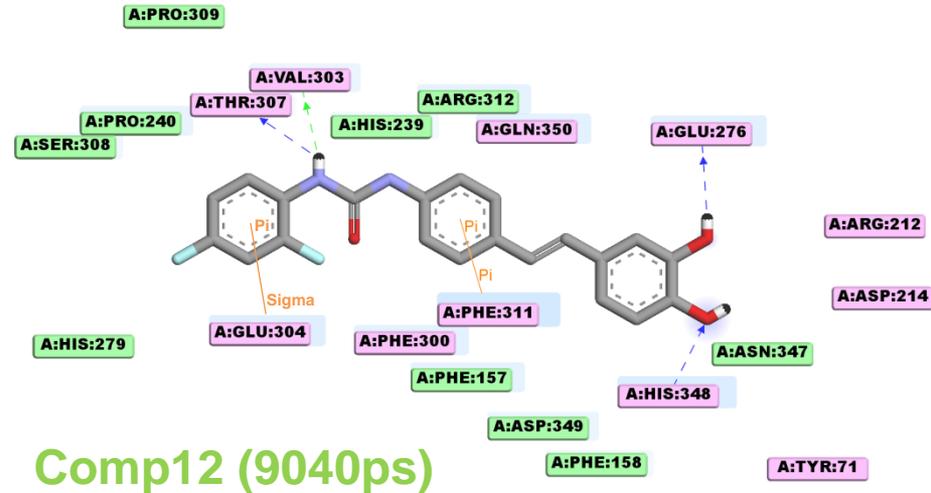
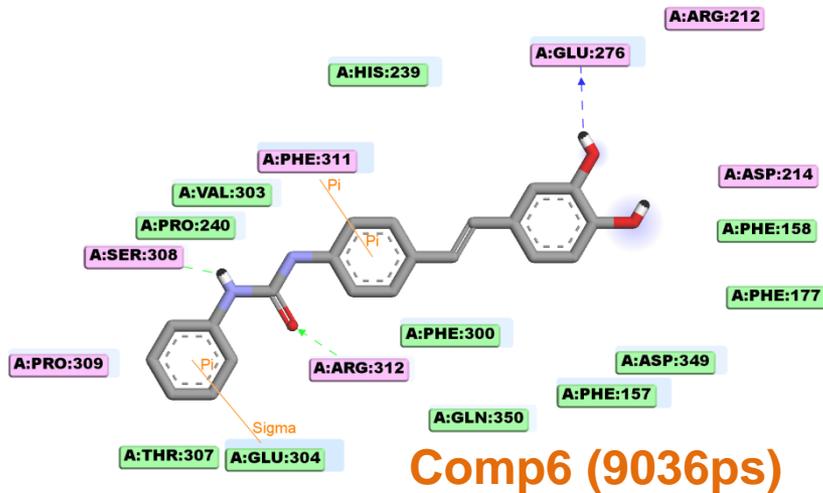
System	Experimental Ki ( $\mu\text{M}$ )	Interaction Energy (kJ/mol)
Comp6	10.6	-245.265
Comp7	10.5	-252.126
Comp11	7.2	-281.353
Comp12	3.2	-287.378
Comp13	4.6	-292.982
Comp14	17	-226.249
Comp16	9.4	-233.907

Correlation graph of **exp. Ki** values with **interaction energy** after 1.5 ns MDS



“After the 1.5 ns MDS, **exp. Ki** was well correlated with **interaction energy**”

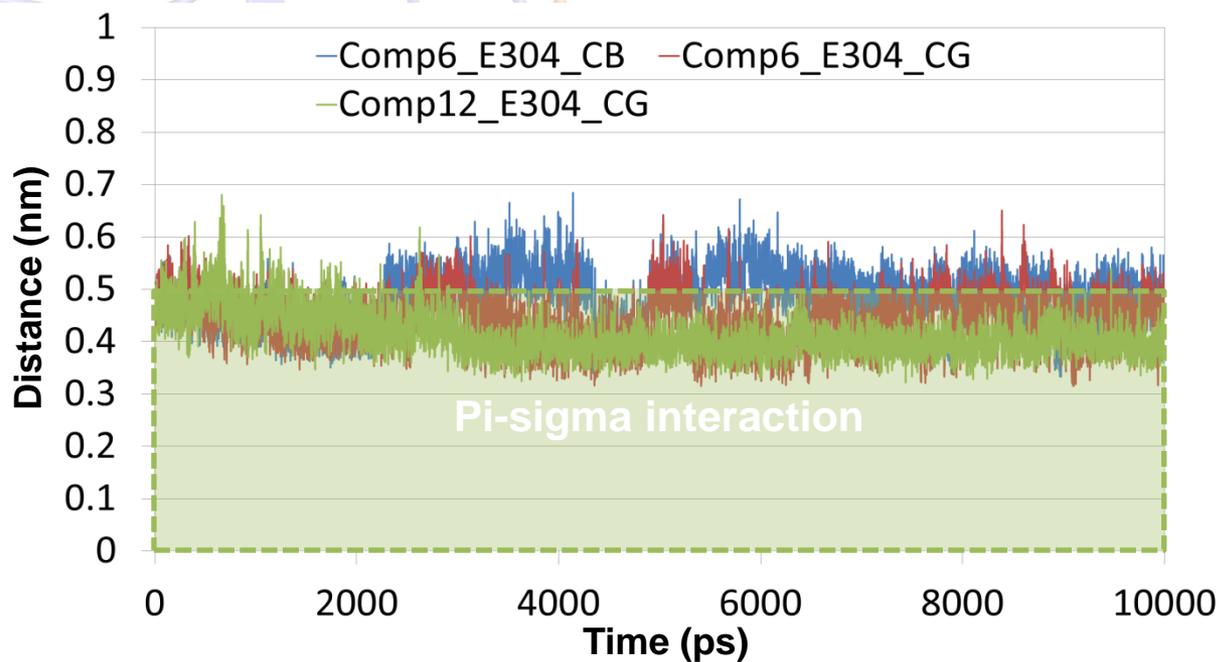
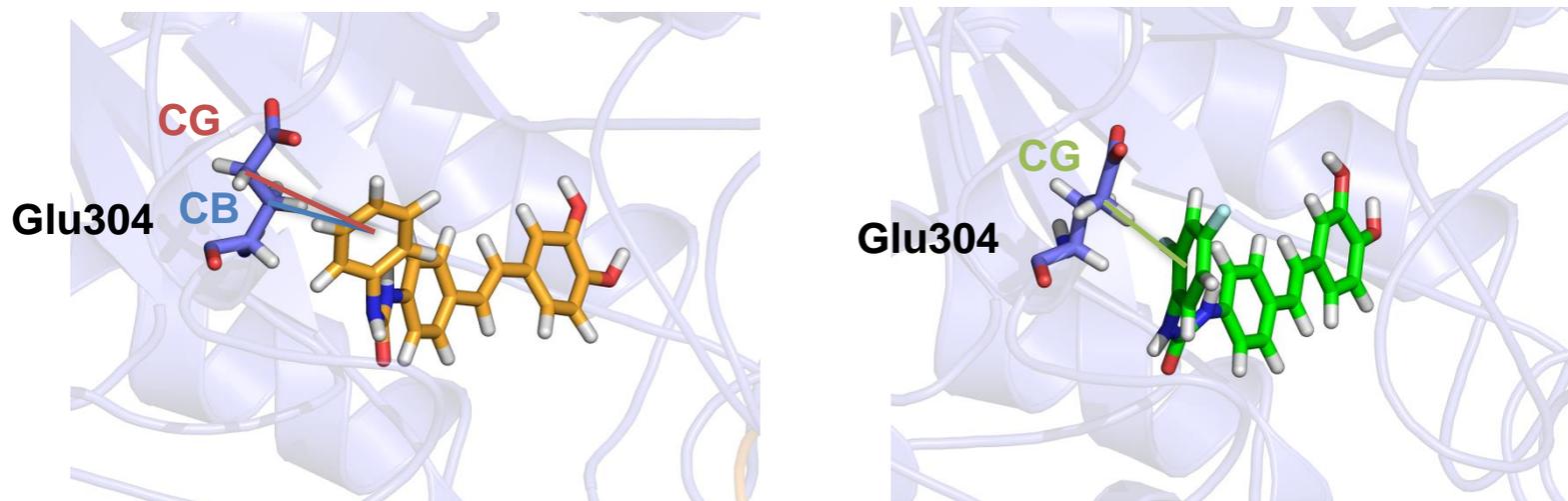
# Binding mode comparison: 10ns MDS representative structures for **Comp6** and **Comp12** systems



Ligand	Protein-ligand interactions		
	Hydrogen bonding residues (<math><3 \text{ \AA}</math>)	Residues involved in charge or polar interaction	Hydrophobic contacting residues
<b>Compound 6</b>	<b>Glu276</b> , Ser308, Arg312	<b>Arg212</b> , <b>Asp214</b> , Pro309, <b>Phe311</b>	<b>Phe157</b> , <b>Phe158</b> , Phe177, <b>His239</b> , <b>Pro240</b> , Phe300, Val303, Glu304, Thr307, <b>Asp349</b> , Gln350
<b>Compound 12</b>	<b>Glu276</b> , Val303, Thr307, His348	Tyr71, <b>Arg212</b> , <b>Asp214</b> , Glu304, Phe300, <b>Phe311</b> , Gln350	<b>Phe157</b> , <b>Phe158</b> , <b>His239</b> , <b>Pro240</b> , His279, Ser308, Pro309, Arg312, Asn347, <b>Asp349</b>

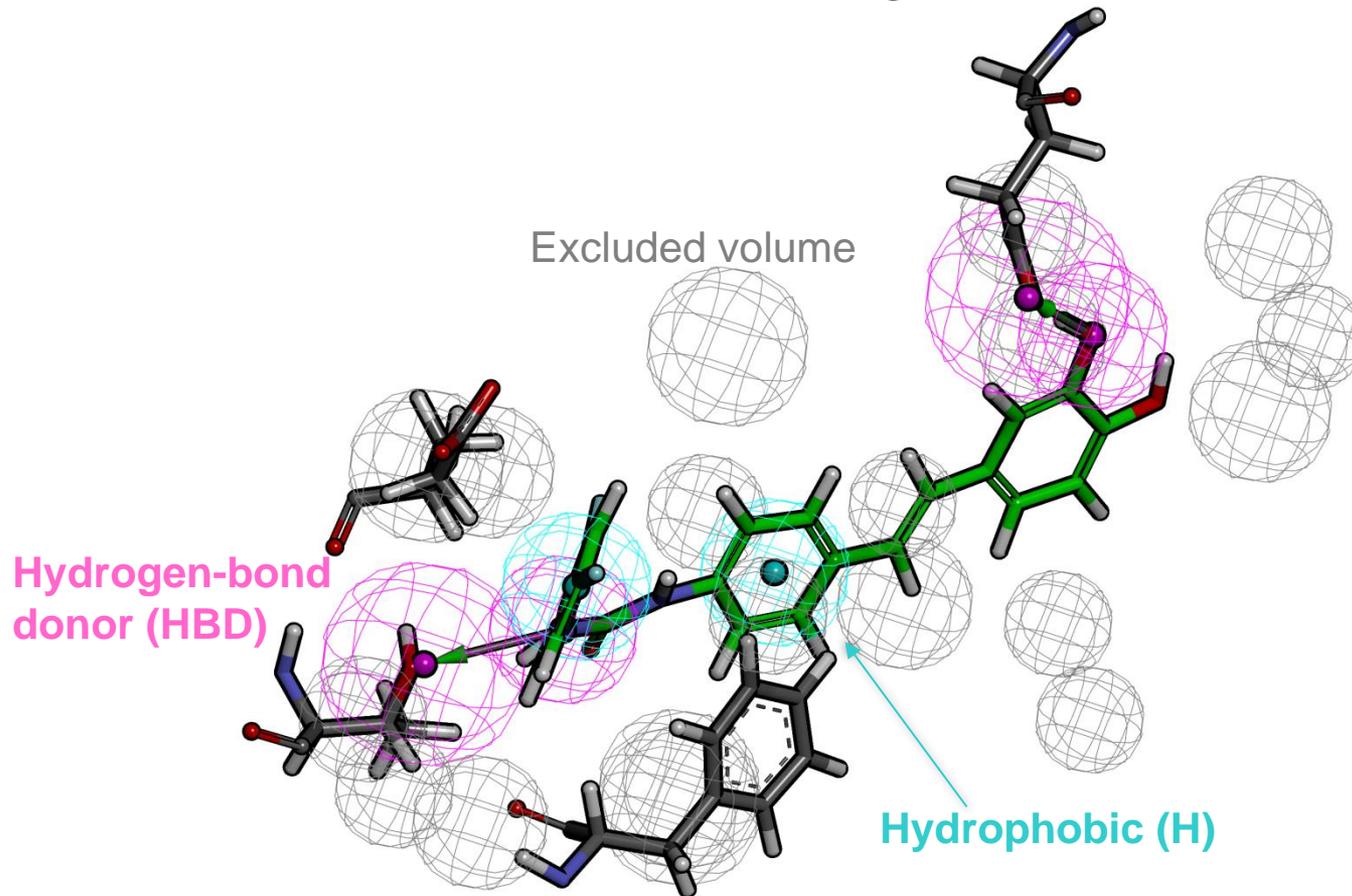
**Glu304 is charged interacting residue in comp12;**  
**But, Glu304 is hydrophobic residue in comp6**

# Pi-sigma interaction comparison



# Generation of Receptor-ligand pharmacophore model

Based on *the final structure*, receptor-ligand *pharmacophore model* was generated for the further virtual screening



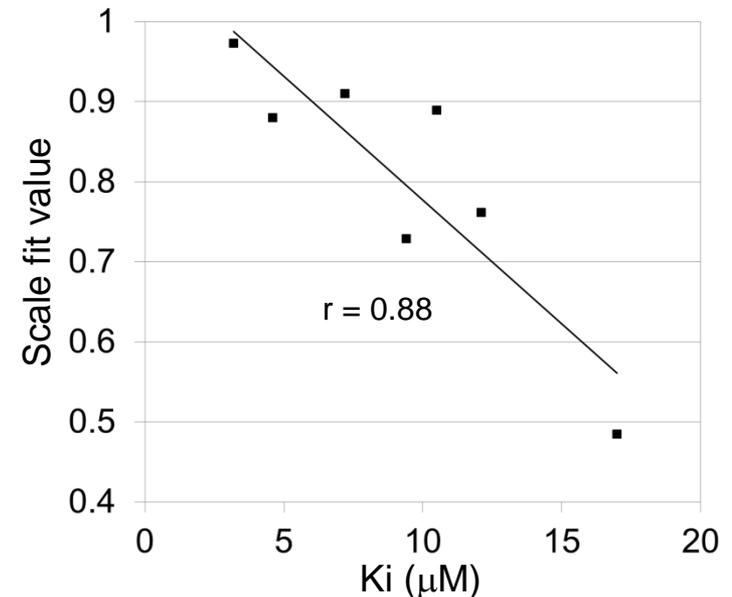
Fit vale: 3.89

# Validation of generated pharmacophore model showing linearity between **exp. Ki** and **scale fit value**

To validate the final binding mode, **exp. Ki** values were correlated with **scale fit value**

System	Time, ns	Exp. Ki ( $\mu\text{M}$ )	Scale fit value
Comp6	10	10.6	-
Comp7	1.5	10.5	0.89
Comp10	1.5	12.1	0.76
Comp11	1.5	7.2	0.91
Comp12	10	3.2	0.97
Comp13	1.5	4.6	0.88
Comp14	1.5	17	0.49
Comp16	1.5	9.4	0.73

Correlation graph of **exp. Ki** values with **scale fit value** after 1.5 ns MDS



“After the 1.5 ns MDS, **exp. Ki** was well correlated with **scale fit value**”

# Combined approach (*molecular docking and MD simulation*) to find out the reasonable binding mode

- © To find out the reasonable binding modes of stilbene urea derivatives within active site of  $\alpha$ -glucosidase

Homology modeling

3D structure of *yeast  $\alpha$ -glucosidase* was constructed using *B. cereus oligo-1,6-glucosidase* (PDB ID: 1UOK)

Molecular docking

*Initial binding modes* of stilbene urea derivatives were found within active site of  $\alpha$ -glucosidase

20 ns MD simulations (Adjustment)

*The lowest energy structure* (adjusted conformation of highest active compound) was selected

Molecular docking

*The final docking poses* within adjusted conformation of *the best structure* were obtained

2 ns MD simulation (Refinement)

*The docked structures* were refined by MDS

Receptor-ligand Pharmacophore generation

Finally, generation of *pharmacophore model* for the further virtual screening

# References

---

## ❖ Books

- Computer simulation of liquids by M. P. Allen & D. J. Tildesley.
- Understanding Molecular Simulation by Daan Frenkel, Berend Smit
- The Art of Molecular Dynamics Simulation by D. C. Rapaport.

## ❖ Popular simulation packages

- GROMACS:
- <http://www.gromacs.org/>
- <http://www.bevanlab.biochem.vt.edu/Pages/Personal/justin/gmx-tutorials/>
- <http://www.gromacs.org/@api/deki/files/198/=gmx-tutorial.pdf>
- <http://md.chem.rug.nl/~mdcourse/index.html>
- CHARMM: <http://www.charmm.org/>
- AMBER: <http://ambermd.org/>
- NAMD: <http://www.ks.uiuc.edu/Research/namd/>

---

**Thank you for your attention~\***

